SOUTHWEST FISHERIES SCIENCE CENTER

NATIONAL MARINE FISHERIES SERVICE SOUTHWEST FISHERIES SCIENCE CENTER P.O. BOX 271 LA JOLLA, CA 92038

APRIL 1994

A MIXTURE MODEL FOR ESTIMATING BYCATCH FROM DATA WITH MANY ZERO OBSERVATIONS: TUNA BYCATCH IN THE EASTERN TROPICAL PACIFIC OCEAN

By

Peter C. Perkins and Elizabeth F. Edwards

ADMINISTRATIVE REPORT LJ-94-07

A mixture model for estimating bycatch from data with many zero observations: Tuna bycatch in the eastern tropical Pacific Ocean

Peter C. Perkins and Elizabeth F. Edwards

Southwest Fisheries Science Center
P.O. Box 271, La Jolla, CA 92038
(619) 546-7017

March 1994

Administrative Report LJ-94-07

Table of Contents

List of Figures

List of Tables

A mixture model for estimating bycatch from data with many zero observations: Tuna bycatch in the eastern tropical Pacific Ocean

Peter C. Perkins
Elizabeth F. Edwards
Southwest Fisheries Science Center
P.O. Box 271, La Jolla, CA 92038
(619) 546-7017

Abstract.

Studies of fisheries bycatch often result in data that are characterized by a smooth distribution of positive amounts of per-set bycatch, but with an extremely large number of zero observations. This discontinuity at zero is difficult to fit with a standard distribution. One approach is to model per-set bycatch with a mixture of two distributions, with one component representing the zero observations and the other representing the observations of positive bycatch. In this report, we describe such a mixture model that is suitable when the bycatch observations have been rounded to integer amounts. In particular, when "rounded" zeros (representing small amounts of bycatch) and "true" zeros (representing no bycatch) are indistinguishable in the data, the mixture model can be used to estimate the proportion of each.

We fit this model to tuna bycatch data collected by observers aboard the U.S. tuna purse-seine fleet in the eastern tropical Pacific Ocean during the years 1989-1992. We use the model to estimate bycatch per set, and allow the model parameters to depend upon one or more covariates. We then show how to estimate mean bycatch per set fishery-wide, by summing out over those covariates. Extensions and limitations are discussed.

1. Introduction.

Many fisheries catch, in addition to the target species, unwanted but unavoidable individuals of other, non-target, species. This "bycatch" of unwanted individuals is generally discarded, and in many fisheries, few, if any, individuals survive their capture and discard. Estimating the extent of such bycatch is becoming increasingly important as fisheries managers more often have to contend with situations where unwanted individuals from a fishery include individuals which are desirable in other contexts. Bycatch in one fishery may include juvenile members of the target species in the same or another fishery, or individuals from threatened, endangered or protected species. For example, juvenile mackerel are caught and discarded, and sea turtles caught and drowned, by shrimp trawlers in the Gulf of Mexico (e.g., Collins and Wenner, 1988; Caillouet et al., 1991). Sharks, billfish, and juvenile tuna are caught by tuna purse-seiners in the eastern tropical Pacific Ocean (e.g., Au, 1991). Numerous tropical fish species are caught by shrimp trawlers around islands in the south Pacific Ocean (e.g., Kulbicki and Wantiez, 1990). Various prohibited species are captured as bycatch in the Bering Sea domestic trawl fisheries (Berger et al., 1989). These are only a few of the numerous fisheries plagued with bycatch.

Although of increasing interest, the amount of bycatch generated by various fisheries remains relatively unstudied. Bycatch is generally discarded without being weighed or measured exactly; few fisheries routinely estimate or measure bycatch in any form. Because complete data are most often lacking, bycatch usually must be estimated rather than reported directly (e.g., Berger et al., 1989). One particularly troublesome problem in bycatch estimation is developing a statistical model that is sufficiently flexible to account for a variety of data types, so that bycatches can be compared from different circumstances.

The U.S. tuna purse-seine fishery in the eastern tropical Pacific Ocean (ETP) presents such a problem, but, in contrast to many other fisheries, also provides an opportunity to develop a solution to the problem, because information on bycatch of tuna (including both non-target tuna species and juveniles of target species) has been collected from the fishery since 1988.

A flexible approach is required to model tuna bycatch from this fishery because the purse-seine vessels capture fish using three distinctly different fishing strategies: "log fishing", "school fishing" and "dolphin fishing". Log fishing is the practice of catching fish by making purse-seine sets on tuna associated with floating objects. These sets usually capture schools of small yellowfin tuna (*Thunnus albacares*) or mixed schools of small yellowfin and like-sized skipjack tuna (*Katsuwonus pelamis*). School fishing is the practice of capturing schools composed purely of (usually small) tuna (again, either pure schools of yellowfin or mixed schools of yellowfin and skipjack), located by surface disturbances created by the schools. Dolphin fishing is the practice of catching tuna located by surface disturbances created by closely associated dolphins (e.g., Orbăch 1977). Tuna associated with dolphins almost always consist of pure schools of large yellowfin. Log fishing generates large amounts of tuna bycatch, very frequently (with almost every set). School fishing generates moderate amounts of tuna bycatch much less often. Dolphin fishing generates small amounts of tuna bycatch, and only very infrequently. Thus, tuna bycatch data from dolphin fishing are characterized by many zero observations, while data from log fishing contain mostly non-zero observations. School fishing presents an intermediate case.

We develop here a method for modelling bycatch per set for these three disparate types of bycatch data, and show how to use the model to estimate mean bycatch per set for each set type. The focus of the present study is development and description of the model as a solution to a common problem in bycatch estimation. Detailed results of applying the model, and its implications for the U.S. tuna purse-seine fishery in the ETP, are the subject of another paper (Edwards and Perkins, in preparation).

We first describe the general characteristics of the available bycatch data, and then fit a statistical model to those data to investigate some of the factors affecting the amount of tuna bycatch per set. We define an appropriately flexible probability distribution to describe the observed data, then use likelihood methods to select relevant covariates and compute parameter estimates for that distribution. Finally, we use those values to estimate mean bycatch per set, both as a function of geographic location and fishery-wide, for each set type. The specific applications and results presented here address only tuna bycatch in the ETP purse-seine fishery, and do not include bycatch of other species (e.g., marine mammals, billfish, sharks, sea turtles). However, the method presented here is generally applicable to any situation involving analysis of data sets characterized by varying proportions of zeros.

2. Data.

Although more detailed reports would be desirable, currently available data for the U.S. purse-seine fleet only include per-set estimates of total tons of tuna discarded. Data on size-classes and species composition of this discarded tonnage are not available, nor are data available on discards of species other than tuna. This report focuses on data from the U.S. fleet only, and does not include direct information about bycatch from the considerable number of non-U.S. purse-seiners fishing in the ETP.

Data were collected by National Marine Fisheries Service (NMFS) or Inter-American Tropical Tuna Commission (IATTC) observers placed aboard U.S. tuna purse-seiners during routine fishing trips to the eastern tropical Pacific Ocean. Each agency provided about 50% of the observers; agencies alternated sending observers on departing trips. Observers recorded time and position of all sets made by U.S. vessels fishing in the ETP during the 31-month "study period" September 1, 1989 - March 30, 1992. Observer coverage was 100% during this period. However, bycatch data were available only from approximately half of the set records during the period September 1, 1989 - July 30, 1990 because during this period only IATTC observers were collecting bycatch data. During the remaining period (August 1, 1990 - March 30 1992) observers from both agencies recorded bycatch for all sets. The beginning of the "study period" is the first time at which both NMFS and IATTC records are available for analysis (prior to this, IATTC data are considered proprietary). The end of the "study period" corresponds to the most recent complete data that were available at the time when this study was initiated.

Although it would have been desirable to take direct measurements of the weight of the bycatch (tons of tuna discarded) for each individual set, this was not feasible. For most sets, observers estimated the bycatch by counting the number of brailers (large fish baskets) used to empty the net after the set, multiplying this number by an estimated tonnage per brailer (based on advice from the fishing captain and other experienced crewmen), and then multiplying this estimated total weight caught by the estimated fraction of non-target tuna in the catch. Observers estimated this fraction of non-target tuna by eye, either by estimating the composition of brailers, or by observing catch sorting on deck. Occasionally the majority of the catch was discarded before being brailed aboard. In these cases, observers estimated discard by first estimating by eye the weight of the total catch, and then estimating the tonnage loaded by brailer. The difference between these two estimates (i.e., tons caught minus tons loaded) was the estimated tons of bycatch and was assumed to include only discarded tuna.

Observer estimates of bycatch tonnage were apparently rounded to integer values, with the rounding interval increasing with the amount of bycatch. For very small amounts of bycatch, weights were rounded to the nearest ton so that it was not possible in these sets to distinguish observations with no bycatch from those with very small amounts (less than one half ton). Increasing difficulty with estimating increasingly large tonnages of bycatch apparently contributed to a systematic tendency in the data towards rounding to the nearest 5 or 10 tons for small and medium estimates of bycatch and to the nearest 25 or 50 tons for the large estimates (Figures 1-5). For sets with moderately small amounts of bycatch, observer estimates tended to be more precise because the bycatch as well as the target fish were brailed aboard the vessel, then sorted on deck where the bycatch could be easily compared to the total catch. For sets with large amounts of bycatch, the fish may not have been brought on board, making precise estimates more difficult and rounding tendencies greater.

The various sources of measurement error and rounding, as described above, introduce into our data uncertainty which we did not attempt to account for. In the absence of data or studies for "ground truthing" observer estimates of bycatch, or a plausible model for the measurement errors, we treated the bycatch weight estimates as exact measurements.

Bycatch was recorded for 59% (2110 of 3590, Table 1) of observed dolphin sets, 76% (960 of 1226) of observed schoolfish sets, and 75% (998 of 1328) of observed log sets. These sets generated 134, 1098, and 9819 tons of observed bycatch, respectively. The relatively small bycatch totals for schoolfish and dolphin sets are due to the large numbers of sets of those two types with zero bycatch recorded. Positive tuna bycatch was recorded in 65% (650 of 998, Table 1) of log sets with bycatch observed, but in only 8% (80 of 960) of schoolfish sets and only 0.7% (10 of 2110) of dolphin sets with bycatch observed. We ignored log and school fishing in area 2 in these analyses, because 0 schoolfish sets and only 10 log sets (4 with estimated bycatch) occurred in this area (Table 1). We also eliminated from the analyses 7 sets in which the entire catch (target catch plus bycatch) was lost due to equipment failure.

TABLE 1. Effort data (numbers of sets) for the U.S. tuna purse-seine fleet fishing in the eastern tropical Pacific Ocean, 1989-1992. Geographic areas as defined in Federal Register (1989). N is the total number of sets in a given area, n is the number of sets with bycatch recorded, and n^+ is the number of sets with positive bycatch recorded.

Set Type	Area	N	n	n^+
Dolphin	1	2496	1445	10
	2	498	272	5
	3	596	393	4
	Total	3590	2110	19
Schoolfish	1	399	279	32
	2	0	0	0
	3	867	681	48
	Total	1266	960	80
Log	1	537	326	257
	2	10	4	4
	3	791	672	393
	Total[1]	1328	998	650

1. Totals for log sets do not include sets in geographic area 2. See text for description.

3. Methods.

3.1 Modelling bycatch per set.

The primary difficulty in defining a probability distribution to model bycatch per set was sufficient flexibility to describe the disparate distributions observed for the three different set types (Figures 1-5). We chose a modified negative binomial distribution, known as the negative binomial with added zeros (Johnson and Kotz, 1969), because this distribution could accommodate the wide range in the proportion of zero observations, as well as the relatively heavy tails in the observed distributions of bycatch for all three set types (see Discussion for two other models that we considered but rejected).

The negative binomial distribution with added zeros is a mixture of a negative binomial distribution and a discrete probability mass at zero. Under this model, bycatch per set is either exactly zero with probability (p) or has a negative binomial distribution with probability (1-p). The negative binomial portion of this distribution can be viewed as representing strictly positive amounts of bycatch, rounded to integer values. Thus, zero values that are part of the negative binomial can be interpreted as observations of small amounts of discard, rounded down to zero. Zero values from the probability mass can be interpreted as exact zeros. The probability function for this modified negative binomial distribution is

$$\Pr\{Y=y\} = \begin{cases} p + (1-p)\left(\dfrac{1}{1+a\mu}\right)^{1/a}, & y=0 \\[2ex] (1-p)\dfrac{\Gamma(y+1/a)}{y!\,\Gamma(1/a)}\left(\dfrac{1}{1+a\mu}\right)^{1/a}\left(\dfrac{a\mu}{1+a\mu}\right)^{y}, & y=1,2,\ldots \end{cases} \qquad [1]$$

where Y is an individual observation (here, a number of tons of bycatch per set for a given set type), (p) is the probability of an observation coming from the "perfect zero" state, (1-p) is the probability of an observation coming from the negative binomial state, and (μ) and (a) are, respectively, the mean and variance parameters of the conditional negative binomial.

The parameter (a) determines the shape of the distribution. As (a) tends to zero, the conditional negative binomial distribution in the mixture tends to a Poisson distribution. As (a) increases, the conditional negative binomial becomes more skewed, with a heavier tail and higher probability of a zero observation. The parameter (p) is a mixing parameter which controls the relative importance of the negative binomial and the probability mass at zero. When (p) is one, the distribution is a probability mass at zero. When (p) is zero, the probability distribution becomes strictly negative binomial and expected bycatch per set is the mean of the negative binomial, (μ).

The expected value for individual observations (i.e., mean tons bycatch per set in this study) from this probability distribution is

$$E[Y] = (1-p)\mu \qquad [2]$$

while the variance of an individual observation is

$$var[Y] = (1-p)(\mu + (a+p)\mu^2). \qquad [3]$$

Geographic area, tons of tuna loaded (i.e., commercial catch), time of day, and month were all factors which we considered using as potential covariates in the analysis. We did not attempt to account for any long-term (i.e., year to year) trend in bycatch rates, because our data included too few years for such an analysis.

We included geographic area in the analysis because of obvious differences in amounts of bycatch and effort between set types and fishing areas in the data. *A priori*, we selected as geographic strata the three areas (Figure 6) currently used to determine comparability of U.S. and non-U.S. dolphin mortality rates (Federal Register, 1989). These roughly define the major fishing areas in the system. The total numbers of sets observed in each area, including sets for which bycatch was not recorded, represent the actual areal distribution of fishing effort during the study period. As shown in Table 1, however, observation of bycatch was not proportional to this true distribution of total effort. In the analysis that follows, it is important to distinguish between the total numbers of sets, denoted by $N_{i,j}$, and the numbers of sets for which bycatch was observed, denoted by $n_{i,j}$. The former define the actual distribution of fishing effort, while the latter simply reflect the sample taken. Because our sample of sets with bycatch recorded was not proportional to the total effort (Table 1), ignoring area in the analysis could lead to biased estimates if the mean bycatch per set differs from area to area for a given set type.

We attempted to include tons of catch loaded in the analysis because it is information that might be available from historical data. We did not, however, find any statistically significant relationship between tons loaded and bycatch per set (Figure 7), and so none of the final parameter estimates depend upon tons loaded.

There are some apparent general relationships between time of day or month and bycatch, but we rejected them as potential covariates for two reasons. First, the sampling was unbalanced, so that some times and months were over-represented and some under-represented. This unbalance resulted from correlations between set type and time of day (e.g., most log sets occurred early in the day), and between area and month (e.g., most dolphin sets occurred in area 2 only during July - September). With such strong imbalances in the data, estimates of coefficients for those covariates would be poor. Second, neither time of day nor month would have contributed to our primary focus in this study, i.e., fishery-wide estimates of mean bycatch for each set type. In contrast to area, our sample of sets for which bycatch was recorded did not appear to be biased with respect to time of day and month. In the absence of an explicit model for effort as a function of those two variables, we assumed that our sample was representative with respect to them. Thus, our estimates of mean bycatch, averaged over time of day and month, would be the same whether or not those covariates were included. Also, the ultimate goal of this study is to predict average annual bycatch (Edwards and Perkins, manuscript in prep.), rather than to model bycatch per set in detail as a function of all possible predictors. Including such relatively uninfluential covariates would increase prediction error as the available information in the data would have to be allocated among more estimated parameters.

We computed estimates of the parameters (p), (μ), and (a) by fitting the model [1] separately to data for each of the three set types, using geographic area as a covariate. To determine an appropriate dependence upon area for each of the three model parameters, we made initial fits for each set type using no areal stratification. We then used stepwise likelihood ratio tests to select or reject more complicated models that included areal dependence for progressively larger numbers of model parameters. At each step, we estimated parameter values by maximizing the likelihood of the observed data for each set type under the modified negative binomial [1], using a quasi-Newton numerical optimization algorithm. We generated all possible models in this procedure and selected the simplest model that could not be significantly improved by adding more terms. It should be noted that because this is not a linear model, significance levels (i.e. p-values) from these likelihood ratio tests are approximate.

We used analytic formulæ to compute standard errors for our estimates of the model parameters (p), (a), and (μ). These formulæ are based on the large-sample normal approximation for the maximum likelihood estimates of the parameters. For comparison, we also computed bootstrap standard errors; see Discussion for a description of the differences.

3.2 Estimating mean bycatch per set.

We used Equation [2] and the maximum likelihood estimates for (p) and (μ) from the best-fit models to estimate mean bycatch per set for each set type. When the fitted models indicated significant areal differences in parameter values (i.e., for school and log sets, see Results), we calculated estimates of mean bycatch per set for each individual area. We then calculated a "pooled" estimate of mean bycatch per set as the weighted average of the area-specific estimates, where weightings were proportional to total effort (number of sets including sets for which bycatch was not recorded) in each area. Where no areal stratification was appropriate (i.e., for dolphin sets, see Results), we calculated only one "pooled" estimate from [2] using the likelihood estimate values for (p) and (μ) derived from the unstratified model.

For example, mean bycatch per set type (i) in area (j) would be estimated as

$$\hat{E}[Y_{i,j}] = (1-\hat{p}_{i,j})\hat{\mu}_{i,j} \qquad [4]$$

and the "pooled" estimate for all areas combined would be estimated as

$$\hat{E}[Y_i]_{pooled} = \sum_j N_{i,j}\hat{E}[Y_{i,j}] \,/\, \sum_j N_{i,j} \qquad [5]$$

where $N_{i,j}$ is the total effort (in number of sets) of type (i) occurring in area (j). Note that this "pooled" calculation is based on the proportion of total sets (including those for which bycatch was not recorded) observed in each area. This is an estimate of the mean bycatch per set over the entire fishery during the study period, but is also valid as a prediction of future bycatch under the assumption that the proportion of effort (sets) in each area remains constant as the actual number of sets varies.

While Equation [4] provides a straightforward way to compute the MLE for the product (1-p)μ, it does require that separate estimates for (p) and (μ) first be computed numerically. More importantly, the variance of the product of (1-$\hat{p}$) and ($\hat{\mu}$) can be difficult to estimate accurately. It is possible, however, to use the likelihood equations for the negative binomial with added zeros to derive simplified forms for the MLE of mean bycatch per set. Specifically, only the product (1-p)μ need be estimated, and simple closed form expressions that do not involve the individual parameter estimates can be derived through algebraic manipulation of the likelihood equations. By the invariance properties of maximum likelihood estimates, these simpler forms give results that are identical to using [4].

With this closed-form approach, one can show that, with no areal stratification, the MLE for the product (1-p)μ is simply the sample mean,

$$\hat{E}[Y] = \bar{y} = (1/n)\sum_k y_k, \qquad [6]$$

where the set type subscript (i) is suppressed for clarity. Similarly, with complete areal stratification, the MLE for each area reduces to the sample mean in that area, and the "pooled" estimate is computed using Equation [5]. In both of these cases, the variance for the MLE of $(1-p)\mu$ can be estimated very easily, using the sample variance of the data.

When only the mixing probability (p) depends on area (i.e., for schoolfish sets; see Results), the MLE for mean bycatch per set in area (j) is slightly more complicated, and reduces to

$$\hat{E}[Y_j] = (n_j^+/n_j) \sum_k y_k^+/n^+, \qquad [7]$$

where n_j^+ and n_j are the number of positive observations and the total number of observations in area (j), y_k^+ are the positive observations in all areas, and n^+ is the total number of positive observations in all areas. Similarly, when only the negative binomial mean (μ) depends on area, the MLE for mean bycatch per set in area (j) reduces to

$$\hat{E}[Y_j] = (n^+/n) \sum_k y_{j,k}^+/n_j^+, \qquad [8]$$

where n^+ and n are the number of positive observations and the total number of observations in all areas, $y_{j,k}^+$ are the positive observations in area (j), and n_j^+ is the total number of positive observations in area (j). Again, [5] is used to compute "pooled" estimates in these latter two cases. Note that the estimates here for different areas are not independent, since both formulæ [7] and [8] involve observations from all areas. In particular, the first term in [7] is an area-specific estimate of the probability of a positive observation, while the second term is a "pooled" estimate of the mean for positive observations. This is consistent with the areal stratification on which [7] is based, and provides more precise estimates of E[Y] than simply taking the sample mean in each area. A similar observation may also be made about [8].

While variance estimates for [6] are easy to obtain, there is no simple analytic result for estimating the variance of [7] or [8] (see Discussion for details). However, although analytic methods could not be applied effectively for all three set types, bootstrap methods can be easily applied in all cases using the previous formulæ for the simplified MLEs of mean bycatch per set. Thus, we used bootstrap methods rather than analytic methods to estimate variances for our estimates of mean bycatch per set.

Bootstrap variance estimates of the estimated mean bycatches per set were computed by repeatedly (B = 1000 times) sampling with replacement from the observed data, computing bootstrap replicate values of the appropriate estimators, and calculating empirical estimates of variance. The resampling procedure varied slightly for each set type, depending upon the particular areal stratification chosen for the model parameters. When no areal stratification was appropriate, data were resampled across all areas. When areal stratification was important, data were resampled by area in the same proportions as the original observations.

Specifically, at each iteration the bootstrap procedure either drew n_i values from the total of n_i bycatch observations for set type (i), or drew $n_{i,j}$ values from the total of $n_{i,j}$ observations for set type (i) in area (j) (see Table 1). From each replicate sample, a bootstrap value, $\tilde{y}^*$, was computed using one of Equations [6], [7], or [8], as appropriate, and then applying Equation [5] if required. For example, in the case of complete areal stratification of the model parameters, the sample mean in area (j) for each replicate sample is, from [6],

$$\tilde{y}_j^* = \sum_k y_{j,k}^* / n_j \qquad [6^*]$$

where * indicates a bootstrap selection. We then use [5] to compute the bootstrap replicate value as the weighted average of the area-specific sample means,

$$\tilde{y}^* = \sum_j [N_j \, \tilde{y}_j^*] \, / \sum_j N_j \qquad [5^*]$$

where N_j is total number of sets in area (j) (including sets without bycatch information recorded). Finally, the bootstrap variance for the estimate of mean bycatch per set is simply the sample variance of the (B) bootstrap replicates, i.e.,

$$\hat{\mathrm{var}}(\hat{E}[Y]) = [\sum_b (\tilde{y}_b^* - (\sum_b \tilde{y}_b^*)/B)^2] \, / \, (B-1) \qquad [9]$$

This procedure was repeated for each of the three set types.

4. Results.

4.1 Modelling bycatch per set.

Variations in data characteristics led to different models for the three different set types. Specifically, we selected different levels of areal stratification in the parameter estimates based on the results of likelihood ratio tests. Geographic area was a statistically significant predictor of bycatch per set for only two of the three set types (log and schoolfish sets).

Positive bycatch from dolphin sets occurred so infrequently that separate models for each geographic area were not statistically tenable. Geographic area, when included as a covariate, failed to produce a significant improvement in the fit, and so for this set type, we selected the model with no areal dependence for any of the parameters. Thus, the estimates for each of the three parameters (p), (a), and (μ) in this case are fishery-wide values (Table 2). The standard error of the mixing parameter (p) for the dolphin model is very small, reflecting the high mixing probability dictated by the extremely large number of zero observations of bycatch. The standard errors of the parameters for the negative binomial portion of the probability distribution ((a) and (μ)) are quite large, reflecting the few positive(19 out of 2110 sets, Table 1) data available for their determination (see Discussion).

At the other extreme, positive bycatch observations were frequent enough for log sets that completely separate, statistically independent models could be developed for geographic areas 1 and 3. In this case, the estimated probability distributions effectively collapsed to an unmodified negative binomial in both areas. Specifically, the numerical optimization failed to converge to a positive value for (p), producing a maximum likelihood estimate of zero for (p) in both areas (Table 2). Using fishing area as a covariate for both the mean (u) and the shape (a) parameters improved the fit significantly (p-value less than 0.001) over simpler models pooling across areas. Because positive observations were so abundant, estimated standard errors for the mean and shape parameters (Table 2) were quite small (see Discussion).

Bycatch from schoolfish sets presented an intermediate case, in which we selected a model which included marginally different maximum likelihood estimates for the mixing probability (p) in areas 1 and 3, but no geographic stratification for the other two parameters, (a) and (μ) (Table 2). There were considerably fewer (80) positive observations for schoolfish sets than for logfish sets, making precise parameter estimation much less likely. Likelihood ratio tests indicated that fishing area should be included as a covariate for either the shape parameter (a) or the mixing probability (p). The approximate p-value for adding areal dependence to the shape parameter was 0.04, while that for the mixing probability was 0.12. These two parameters are similar in the effect they have on the estimated distribution. Increasing either one increases the probability of a zero observation, although increasing (a) also increases the probability of a large observation. Including areal dependence for both parameters simultaneously, or for the mean, did not significantly improve the fit. We chose to include areal dependence only for the mixing probability for two reasons. First, the small number of positive observations for schoolfish sets limits the precision of the shape estimate (see Discussion). Second, the difference in the estimated shape between areas was largely due to two unusually large observations in area 3. Without these two observations, the difference in estimated shapes was reduced, and the significance levels of the two different models were nearly equal (approximate p-values of 0.09). As was the case for dolphin sets, the predominance of zeros in the schoolfish bycatch data set led to small estimated standard errors for the mixing probability (p) but to large estimated standard errors for the mean and shape parameters.

The estimates for the mixing probability parameter (p) for the three set types imply that essentially all dolphin sets (98%) involve no bycatch at all, while log sets always involve some bycatch, although frequently in small amounts. This conclusion is based on the interpretation of zeros derived from the two components of the probability model which we fit (see Discussion for more implications of this interpretation). Observer experience[1] indicates that this result is consistent with observed patterns for dolphin and log sets.

The estimated shape parameters varied widely for the three set types (Table 2). However, because of the large standard error estimates for the schoolfish and dolphin shape parameters, it is not possible to make any strong statements regarding shape as a function of set type from these data. As mentioned above, the estimated shape parameter for schoolfish sets was strongly affected by the presence of two unusually large observations (100 and 125 tons of bycatch) in area 3. Repeating the analysis without these two observations led to a shape estimate of 3.75 (s.e. = 2.10), more similar to those for log sets.

TABLE 2. Maximum likelihood parameter estimates for the negative binomial with added zeros fit to tuna bycatch data from the U.S. tuna purse-seine fleet fishing in the eastern tropical Pacific Ocean, 1989-1992. See text for a description of the parameters and geographic areas. Estimates of standard error appear in parentheses.

	Dolphin Sets	Log Sets		Schoolfish Sets	
		Area 1	Area 3	Area 1	Area 3
p	.982 (.015)	0 (0)	0 (0)	.715 (.182)	.825 (.111)
a	3.87 (5.36)	2.34 (0.19)	3.93 (0.25)	7.20 (6.28)	
μ	3.53 (3.21)	15.4 (1.3)	7.09 (0.55)	5.53 (3.60)	

4.2 Estimating mean bycatch per set.

Since the model we fit for log set bycatch reduced to a simple negative binomial distribution (with $\hat{p} = 0$), the estimates of mean bycatch per log set in each fishing area are just the corresponding mean parameters (μ_j). Mean bycatch per schoolfish or dolphin set was estimated using Equation [4].

Estimates of mean bycatch per log set were an order of magnitude larger than for schoolfish sets and two orders of magnitude higher than for dolphin sets (Figure 8). Most of this difference is due to the wide range in the estimated proportion of sets with zero bycatch. By comparison (Table 2), estimated mean parameters for the negative binomial component of the model differ by less than a factor of five. Thus, the model that we fit indicates that on average, there is a considerable difference among set types in per-set bycatch, although for sets in which bycatch actually occurs, there is comparatively less difference in the amount.

Mean bycatch for log sets was estimated at 10.5 tons per set pooled over areas, ranging from 7.1 tons per set in area 1 to more than double that value (15.4 tons per set) in area 3 (Figure 8). Mean bycatch for schoolfish sets was estimated at 1.16 tons per set pooled over areas, ranging from 1.57 tons in area 1 to 0.97 tons in area 3. Mean bycatch per set for dolphin sets was estimated at 0.06 tons per set fishery-wide. Implications of these results for the fishery are discussed in another study (Edwards and Perkins, in prep.).

1. personal communication, Al Jackson, SWFSC

The coefficients of variation (c.v.'s) for the estimates of mean bycatch per schoolfish and dolphin sets (21% and 33%, respectively) are dramatically smaller than those for the individual parameter estimates of (a) and (μ) (Figure 9). This is because estimating mean bycatch per set (i.e., (1-p)μ) is a more robust procedure than estimating the individual parameters; see Discussion for a full description. In the case of log sets, the c.v.'s for the estimates of E[Y] and (μ) differ (Figure 9), even though in this case, the model reduced to a negative binomial distribution where E[Y] = μ. The c.v.'s differ because, in estimating variances for the individual parameter estimates, we used analytic approximations, while in estimating variances for mean bycatch, we used bootstrap methods (see Methods). For a more detailed description of the differences, see Discussion.

Note that the fishery-wide estimates for log and schoolfish sets are not simply the average of the estimates in each fishing area. This is because the number of sets in each area for which bycatch was recorded was not proportional to the actual number of sets made in that area. This imbalance was an important reason for including geographic area in the analysis. Non-proportional sampling was not a factor for dolphin sets, as the estimated bycatch in that case was the same for all fishing areas.

5. Discussion.

5.1 Estimating model parameters and mean bycatch.

It can be shown from the likelihood equations for the negative binomial with added zeros that estimates for the parameters (a) and (μ) depend solely on the positive observations in the data. The estimate for the parameter (p) depends on all the data, but is strongly dependent on the proportion of zero observations. Thus, the precision of the estimates for (a) and (μ) can be very poor if the data contain few positive observations, while the precision of the estimate for (p) may still be very good. This was the case in our model parameter estimates for dolphin sets, and, to a slightly lesser degree, schoolfish sets. This same observation also applies to the estimate of mean bycatch per set, (1-p)μ. The estimate of this product is more robust than estimates of the individual parameters involved in it, because it does not depend solely on either the positive observations or the proportion of zeros.

5.2 Estimating variances for model parameter estimates.

The analytic approximation formulæ that we used to estimate the variance of the individual parameter estimates are based on the asymptotic normality of maximum likelihood estimates. Specifically, for each set type, we derived the expected Fisher information matrix as a function of the model parameters (p), (a), and (μ), and then calculated the inverse of this matrix evaluated at the MLEs of those parameters. We then took the diagonal elements of this inverse matrix as estimates of the variances of the parameter estimates. Since this method uses the MLEs for (p), (a), and (μ) (rather than using their unknown "true" values) in the information matrix, it suffers from the well-known (e.g., Efron, 1992) but unavoidable tendency for ML estimates of variance to be biased downwards. We did not attempt to "bias correct" these variance estimates.

In addition to the analytic formulæ described above, we attempted to use bootstrap methods to estimate variances for the model parameters (p), (a), and (μ). Bootstrapping can be more robust in that it does not require any assumptions beyond the sample being representative of the underlying process. However, we could not use the method effectively for dolphin sets because there were so few sets observed with positive bycatch recorded (19 out of 2110 sets for which bycatch was recorded; Table 1). In resampling for the bootstrap, approximately one third of the samples contained too few positive observations for the maximum likelihood algorithm to converge. However, results from bootstrapping can be used to shed light on the validity of the normal approximation implicit in the standard errors which we report.

Histograms of dolphin set parameter estimates for the bootstrap samples that did converge were very skewed (Figure 10). By implication, the normal-approximation variance estimates for the dolphin data, while convenient, are probably not very satisfactory. For schoolfish data, bootstrap estimates of standard error were consistently higher than the analytic approximations, indicating that the latter may be optimistic. Histograms of the bootstrap replicate parameter estimates in this case were slightly skewed, due to a small number of unusually large observations. For log sets, bootstrap standard errors were very similar to those from the analytic approximations, and the corresponding histograms were close to normality. The analytic estimates in this case are probably appropriate.

One alternative to bootstrapping in this case might be the use of likelihood intervals. This method provides an analytic means of estimating the precision of parameter estimates when the assumptions required for the information matrix approach are questionable. These intervals do not assume normality of the MLEs, and are not necessarily symmetric about the estimates. However, for large numbers of parameters, calculating likelihood intervals can be computationally difficult.

5.3 Estimating variances for mean bycatch per set estimates.

In an attempt to derive analytic formulæ for the variance of our estimates of E[Y], we manipulated the likelihood equations for the negative binomial with added zeros and found simplified forms for the MLE of E[Y]. In some cases, the simplified form reduces to the sample mean, Equation [6], and the variance of that estimator is simply (suppressing area and set type subscripts for simplicity)

$$\mathrm{var}(\hat{E}[Y]) = (1/n)E[Y] = (1/n)(\mu + a\mu^2), \qquad [10]$$

which can be estimated by substituting in MLEs for (a) and (μ). More simply, using the fact that the estimator is just the sample mean, the minimum variance unbiased estimate of [10] is the sample variance,

$$\hat{var}(\hat{E}[Y]) = [\sum_i (y_i - \bar{y})^2] / (n-1), \qquad [11]$$

where $\bar{y}$ is the sample mean. In other cases, the simplified forms for the MLE of E[Y] are slightly more complex (Equations [7] and [8]), and Equations [10] and [11] no longer apply. It is possible to derive expressions, analogous to Equation [10], for the variance of Equations [7] and [8] in terms of the three model parameters (p), (a), and (μ). However, these formulæ are so complex as to be of no practical use in estimation, and no expression analogous to Equation [11] seems possible.

Thus, for consistency, we used bootstrap methods in all cases. However, when possible (i.e., log and dolphin sets), we also estimated variances using Equation [11], and found that the two sets of results agreed to within about 5%. Likelihood intervals may also be feasible for estimating precision in this case.

We did not encounter the problem discussed in Section 5.2 (i.e., too few positive observations in the case of dolphin sets) when bootstrapping variance estimates of mean bycatch per set. As discussed in Section 5.1, estimates for mean bycatch per set (1-p)μ do not depend solely on positive observations.

5.4 Rounding errors in the observations.

The model used in this study, the negative binomial with added zeros, is comprised of two probabilistic components. As noted in Section 3.1, zero values derived from the negative binomial component can be interpreted as observations of small amounts of discard, rounded down to zero, while zero values from the probability mass component can be interpreted as exact zeros. This interpretation is based on the assumption of an underlying continuous distribution for positive discard amounts (e.g., a gamma distribution), upon which rounding errors have been superimposed.

One consequence of this interpretation is that the mean amount of bycatch that should be associated with "perfect" zeros is zero tons, while the mean amount that should be associated with "negative binomial" zeros is nonzero. Thus, strict adherence to this interpretation of zeros leads to the conclusion that Equation [4] may be an underestimate of E[Y]. However, if we assume a strictly decreasing underlying distribution for positive bycatch, symmetric rounding of amounts larger than one half ton would tend to increase the estimate. In the absence of a specific model for the rounding errors, we did not attempt to correct for any bias due to rounding.

5.5 An alternative algorithm for maximizing the likelihood.

To maximize the likelihood for the individual model parameters (p), (a), and (μ), we used a quasi-Newton maximization algorithm. An alternative method available for mixture models (e.g., McLachlan and Basford, 1988; Lambert, 1992) uses the EM algorithm to maximize the likelihood. This method is generally applicable to distributions with added zeros. In situations with many covariates for the mixing probability (p) and conditional mean (μ) parameters, it provides an alternative to the high-dimensional gradient search required by standard numerical optimization algorithms. The algorithm can be implemented using standard regression techniques for generalized linear models.

We applied the EM algorithm to the negative binomial with added zeros, using a combination of logistic regression to maximize likelihood for (p) and a quasi-likelihood negative binomial regression for (a) and (μ) (Lawless, 1977). This algorithm assumes a constant shape parameter (a), although it may also be possible to modify the method when (a) depends on one or more categorical or continuous covariates. However, this approach was not successful for the current data set because the logistic regression for the mixing probability failed to converge, since, in the case of log sets, the ML estimate for (p) was zero.

5.6 Alternative models considered.

For this analysis, we used the negative binomial with added zeros to model per-set bycatch. We considered but rejected two alternative models: the Δ-distribution (a mixture of a probability mass at zero with a lognormal (Aitchison, 1955; Pennington, 1983)), and a gamma distribution mixed with a probability mass at zero (Coe and Stern, 1982). Both models have been used in similar cases where the data to be analyzed have contained large numbers of zeros. We rejected these models for this study because both were unsuited to our data. The Δ-distribution assumes that the natural logs of the positive observations are distributed normally, or can be so transformed, and this assumption was not plausible. The data in this analysis were rounded to the nearest ton and the mode of the positive observations was at one ton. Thus, no transformation could bring these data to even approximate normality. The gamma mixture model was not appropriate for the current data because maximum likelihood estimation for a highly skewed gamma distribution depends heavily upon small (near zero) observations. In this study, all observations in that region were rounded to either zero or one, implying a large relative measurement error, and therefore potentially poor accuracy. Another more fundamental reason why we rejected these two models was that both models mix a continuous distribution on the positive numbers with a probability mass at zero, and assume that observations from each component remain distinguishable. In the current data set, small positive observations are grouped together with zero observations, and using a negative binomial in the mixture allows the model to distinguish between "true zeros" (actual absence of bycatch) and "rounded zeros" (bycatch so small that it was ignored or missed).

5.7 Conclusions.

The methods developed here were used to model fisheries bycatch data which were rounded to integer values and which included widely varying numbers of zero observations, depending on one or more covariates. The usual models for integer-valued data, e.g. the Poisson distribution, did not fit the current data at all well because of the extreme skewness of some of the observed distributions. The negative binomial with added zeros is more flexible than the standard models, and provided a much better fit to the current data.

Modelling these data with a parametric probability distribution allowed us to describe patterns in the bycatch in some detail, for example, estimating the percentage of "true zeros" vs. "rounded zeros". In contrast, computing non-parametric estimates of mean and variance would not give any indication of the patterns in the individual observations. While average or total bycatch is of significant interest, it is also important to quantify the amount of bycatch possible for an individual set. Assuming that the parametric model is accepted as appropriate, one can estimate, for example, what the probability is that, due to random chance alone, bycatch from a particular boat will exceed a certain limit in a fixed number of sets.

Modelling the data in a regression framework allowed us to test for dependence of the model parameters upon the covariates. In turn, dependencies of the model parameters were transformed into statements about the mean bycatch per set as a function of set type and geographic area. In particular, we were able to incorporate areal dependence into our estimates of mean bycatch per set only when appropriate. Additionally, we were able to distinguish whether differences in mean bycatch were due to differences in the proportion of zero observations, or due to differences in the distribution of positive observations.

In this report we modelled data from observations of fisheries bycatch, however the model is more generally applicable to integer-valued data which exhibit a large proportion of zero observations combined with long positive tails. Both categorical or continuous covariates may be incorporated into the model.

Acknowledgments.

We thank Al Jackson for his generous help and advice about data collection methods and data form interpretation, and the Inter-American Tropical Tuna Commission for making available their data on U.S. tuna vessel tuna bycatch. We especially thank Cleridy Lennert of the IATTC for her helpful suggestions.

Literature Cited.

Aitchison, J. 1955. On the distribution of a positive random variable having a discrete probability mass at the origin. *JASA* 50: 901-908.

Au, D.W. 1991. Polyspecific nature of tuna schools: shark, dolphin, and seabird associates. *Fish. Bull.* 89(3):343-354.

Berger, J.D., R.F. Kappenman, L.L. Low, and R.J. Marasco. 1989. Procedures for bycatch estimation of prohibited species in the 1989 Bering Sea domestic trawl fisheries. *NOAA Technical Memorandum* NOAA-TM-NMFS-F/NWC-173. National Marine Fisheries Service, Seattle. 23 pp.

Caillouet, C.W., Jr., M.J. Duronslet, A.M. Landry, Jr., D.B. Revera, D.J. Shaver, K.M. Stanley, R.W. Heinly, and E.K. Stabenau. 1991. Sea turtle strandings and shrimp fishing effort in the northwestern Gulf of Mexico, 1986-89. *Fish. Bull.* 89(4):712-718.

Coe, R. and R.D. Stern. 1982. Fitting models to daily rainfall. *J. Appl. Meterol.* 21:1024-1031.

Collins, M.R. and C.A. Wenner. 1988. Occurrence of young-of-the-year king, *Scomberomorus cavalla*, and Spanish, *S. maculatus*, mackerels in commercial-type shrimp trawls along the Atlantic coast of the southeast United States. *Fish. Bull.* 86(2):394-397.

Efron, B. 1992. Six questions raised by the bootstrap, in *Exploring the Limits of the Bootstrap*, LePage, R. and L. Billard, eds. John Wiley and Sons, New York. 426 pp.

Federal Register 1989. Interim final rule with request for comments. Tuesday, 7 March, 1989. Vol. 54(43):9438-9451, Rules and Regulations.

Johnson, N.L. and S. Kotz. 1969. *Distributions in Statistics: Discrete Distributions.* Houghton Mifflin, Boston. 187 pp.

Kulbicki, M. and L. Wantiez. 1990. Comparison between fish bycatch from shrimp trawlnet and visual censuses in St. Vincent Bay, New Caledonia. *Fish. Bull.* 88(4):667-675.

Lambert, D. 1992. Zero inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34(1):1-14.

Lawless, J.F. 1987. Negative binomial and mixed Poisson regression. *Can. J. Stats.* 15(3):209-225.

McLachlan, G.J. and K.E. Basford. 1988. *Mixture Models: Inference and Applications to Clustering.* Marcel Dekker, New York. 253 pp.

Pennington, M. 1983. Efficient estimators of abundance for fish and plankton surveys. *Biometrics* 39(1):281-286.

FIGURE 1. School set discards in area 1. Bars indicate observed frequencies, lines indicate fitted frequencies.

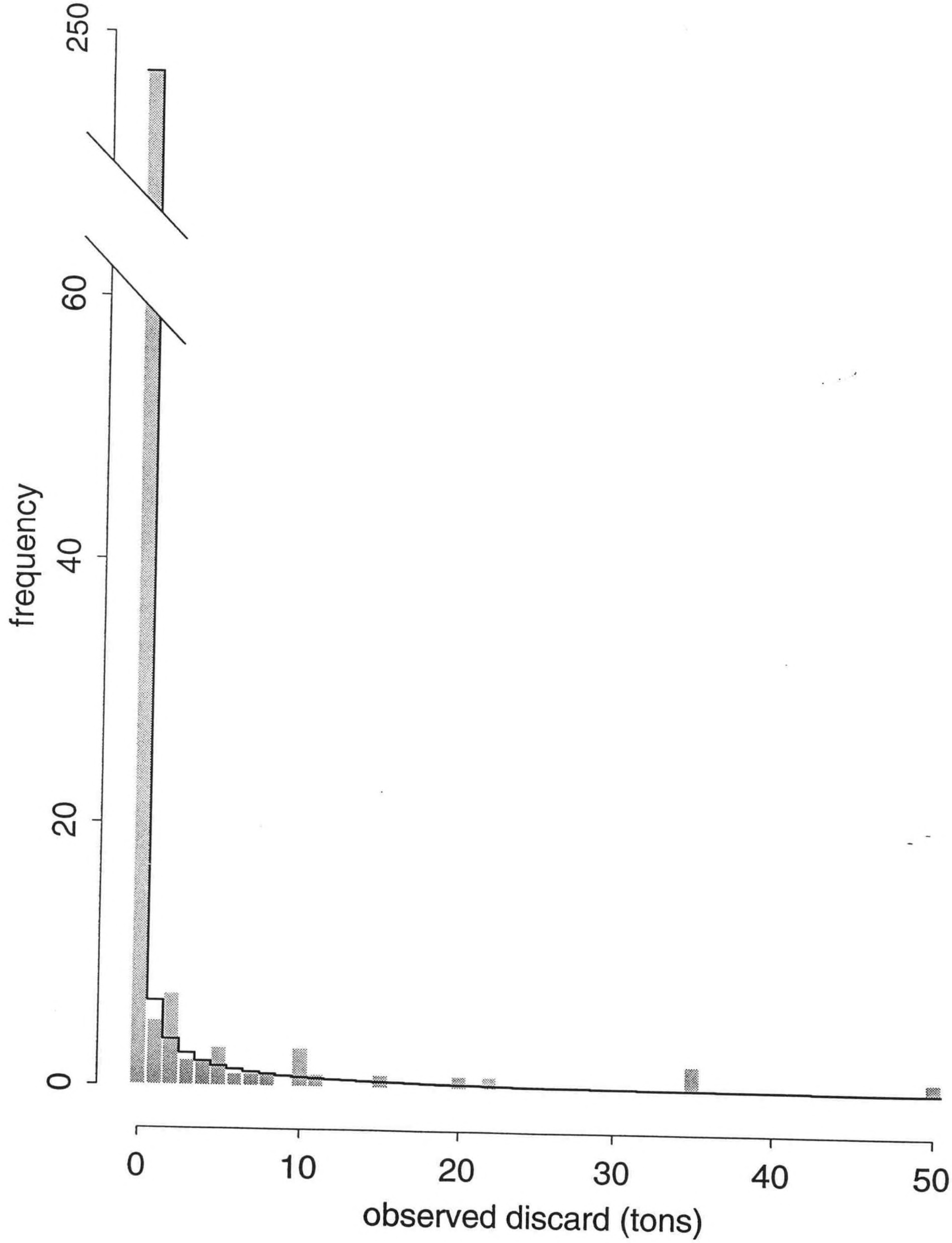

FIGURE 2. School set discards in area 3. Two observations larger than 100 tons not shown. Bars indicate observed frequencies, lines indicate fitted frequencies.

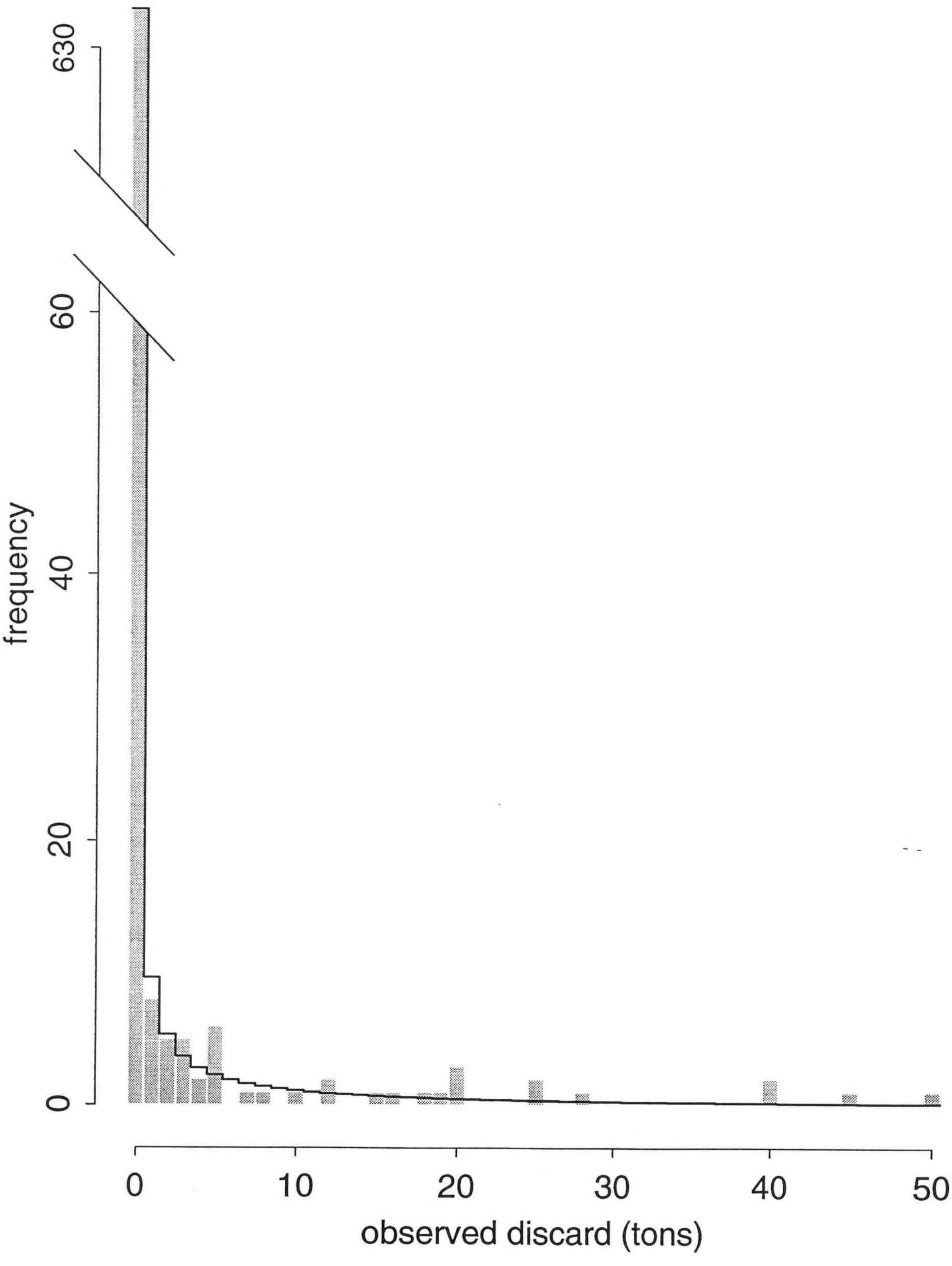

FIGURE 3. Dolphin set discards in all areas. Bars indicate observed frequencies, lines indicate fitted frequencies.

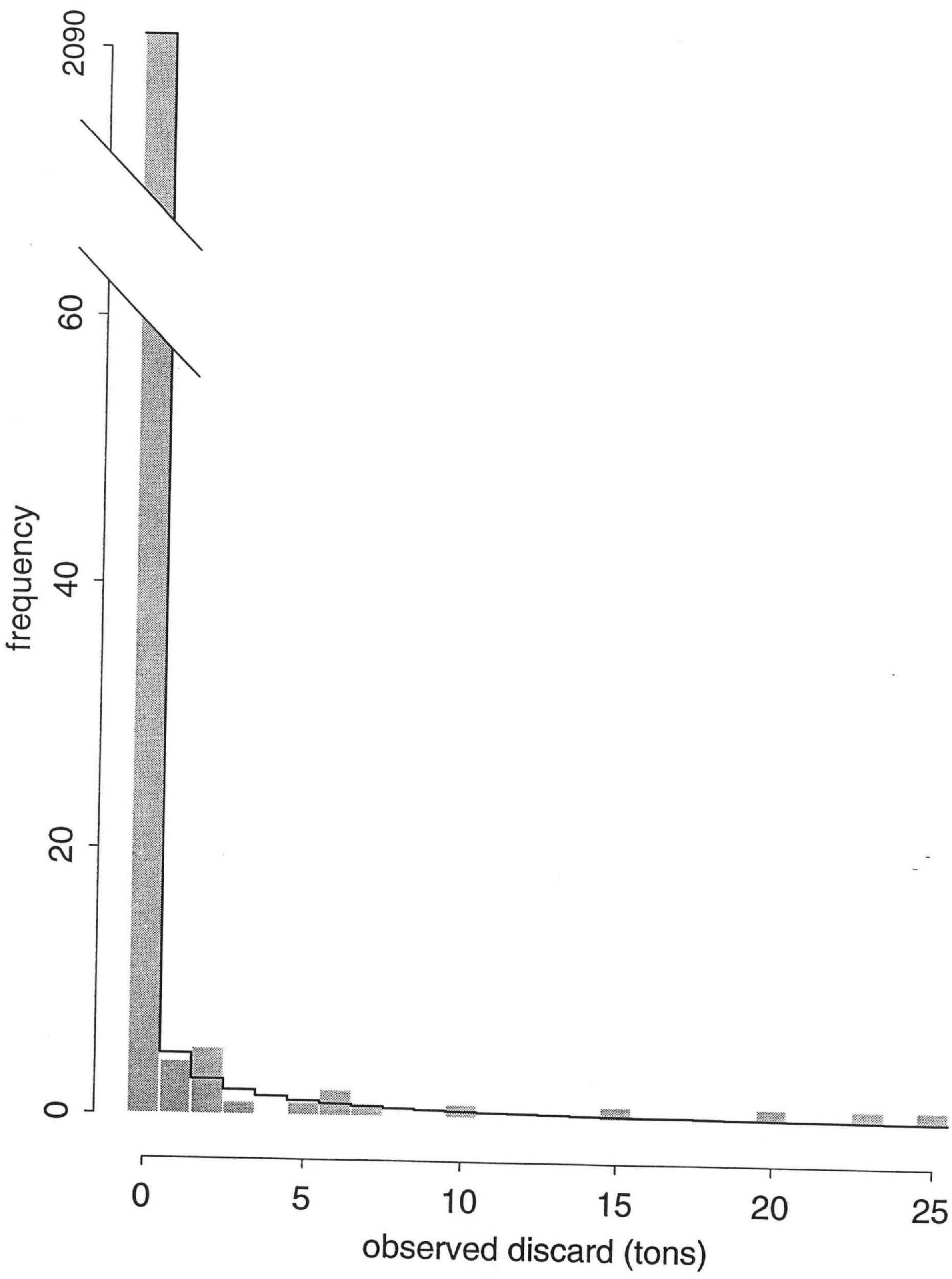

FIGURE 4. Log set discards in area 1. Ten observations larger than 100 tons not shown. Bars indicate observed frequencies, lines indicate fitted frequencies.

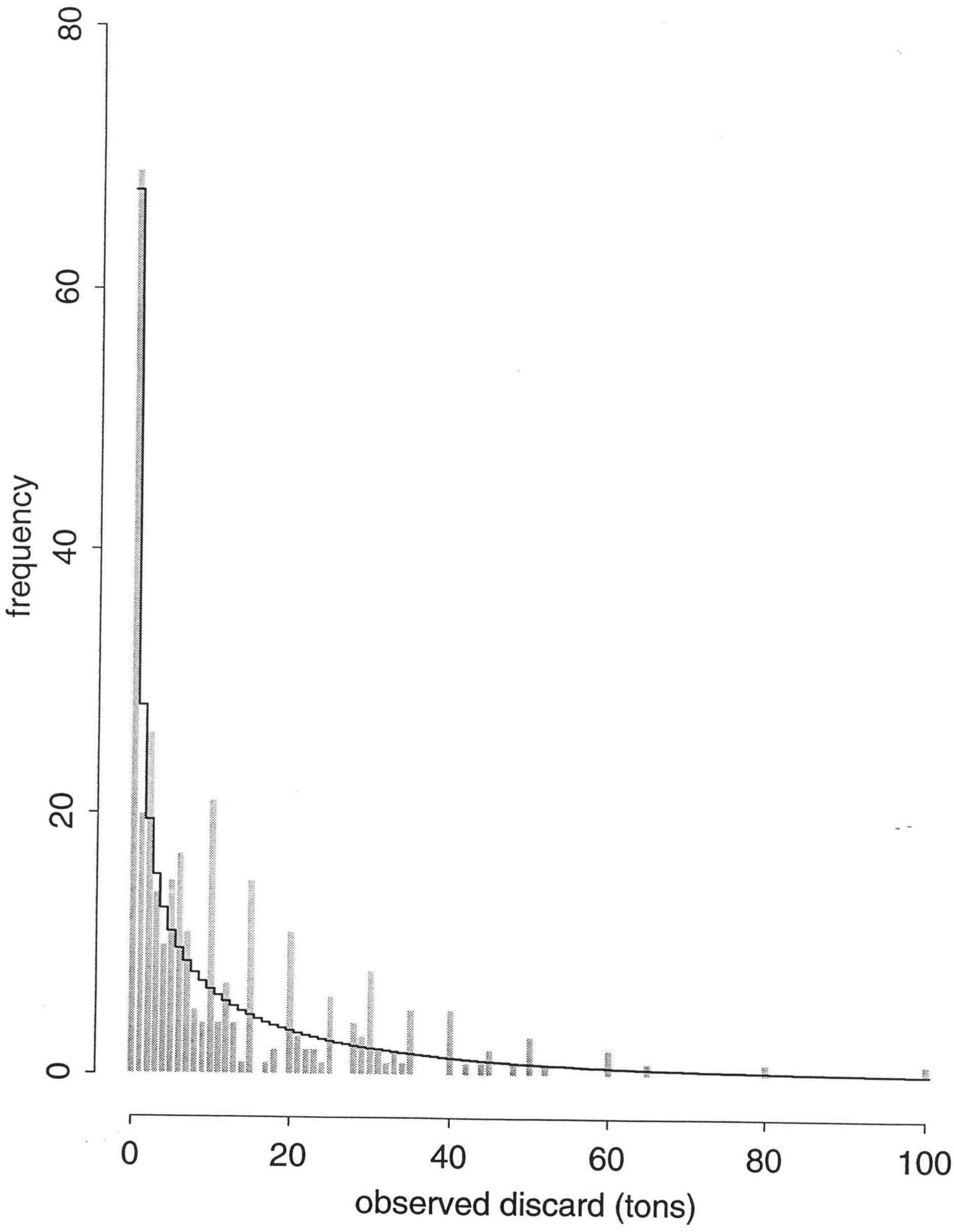

FIGURE 5. Log set discards in area 3. Three observations larger than 100 tons not shown. Bars indicate observed frequencies, lines indicate fitted frequencies.

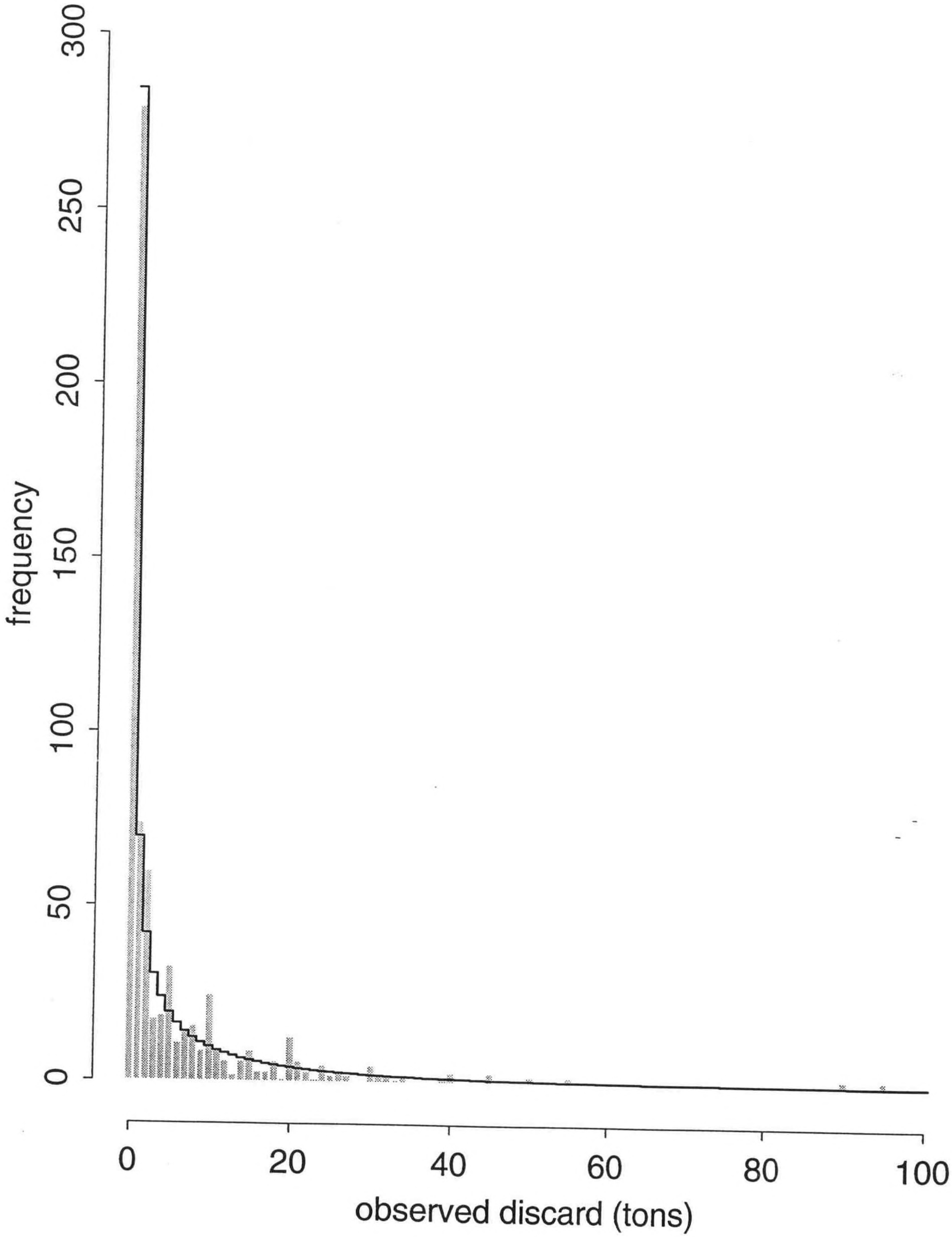

FIGURE 6. Geographic strata used in developing models to estimate mean bycatch per set for the U.S. tuna purse-seine fleet fishing in the eastern tropical Pacific Ocean, 1989-1992 (Federal Register, 1989).

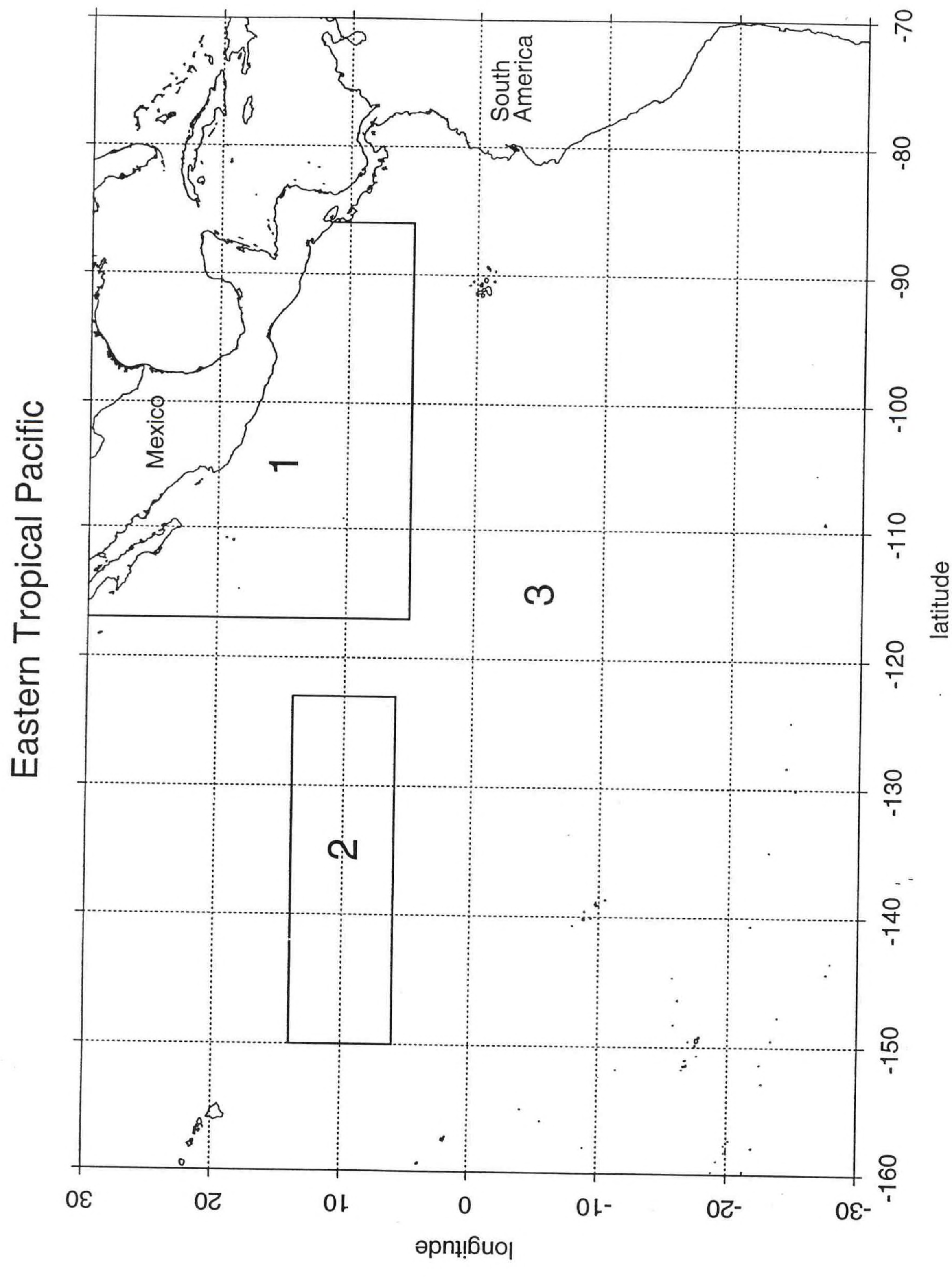

FIGURE 7. (Lack of) relationship between tons of tuna bycatch and tons of tuna loaded for the U.S. tuna purse-seine fleet fishing in the eastern tropical Pacific Ocean, 1989-1992.

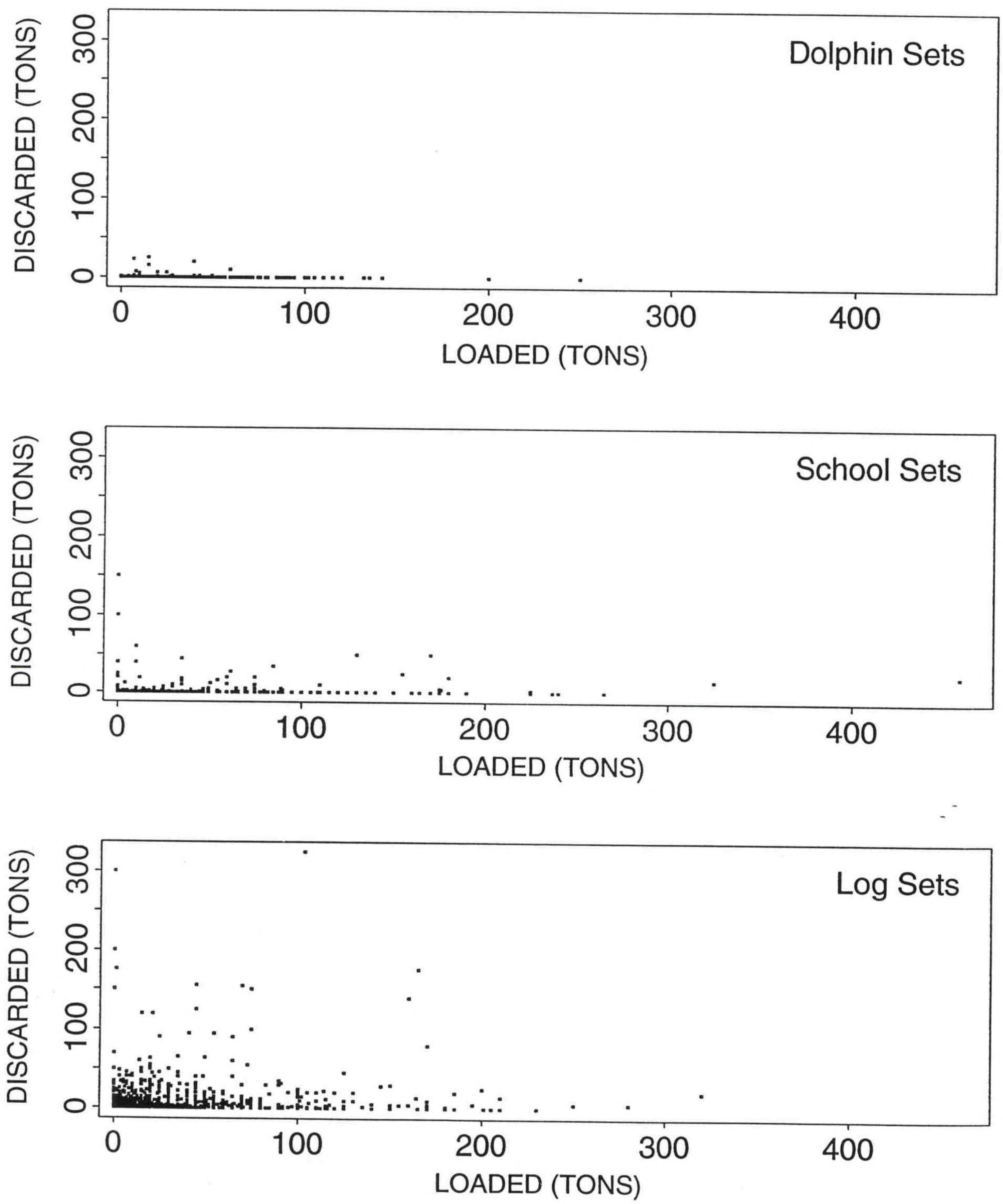

FIGURE 8. Estimated mean tuna bycatch per set for the U.S. tuna purse-seine fleet fishing in the eastern tropical Pacific Ocean, 1989-1992. Geographic areas as defined in Federal Register (1989). Pooled estimates are fishery-wide, across all areas. Standard errors indicated by error bars.

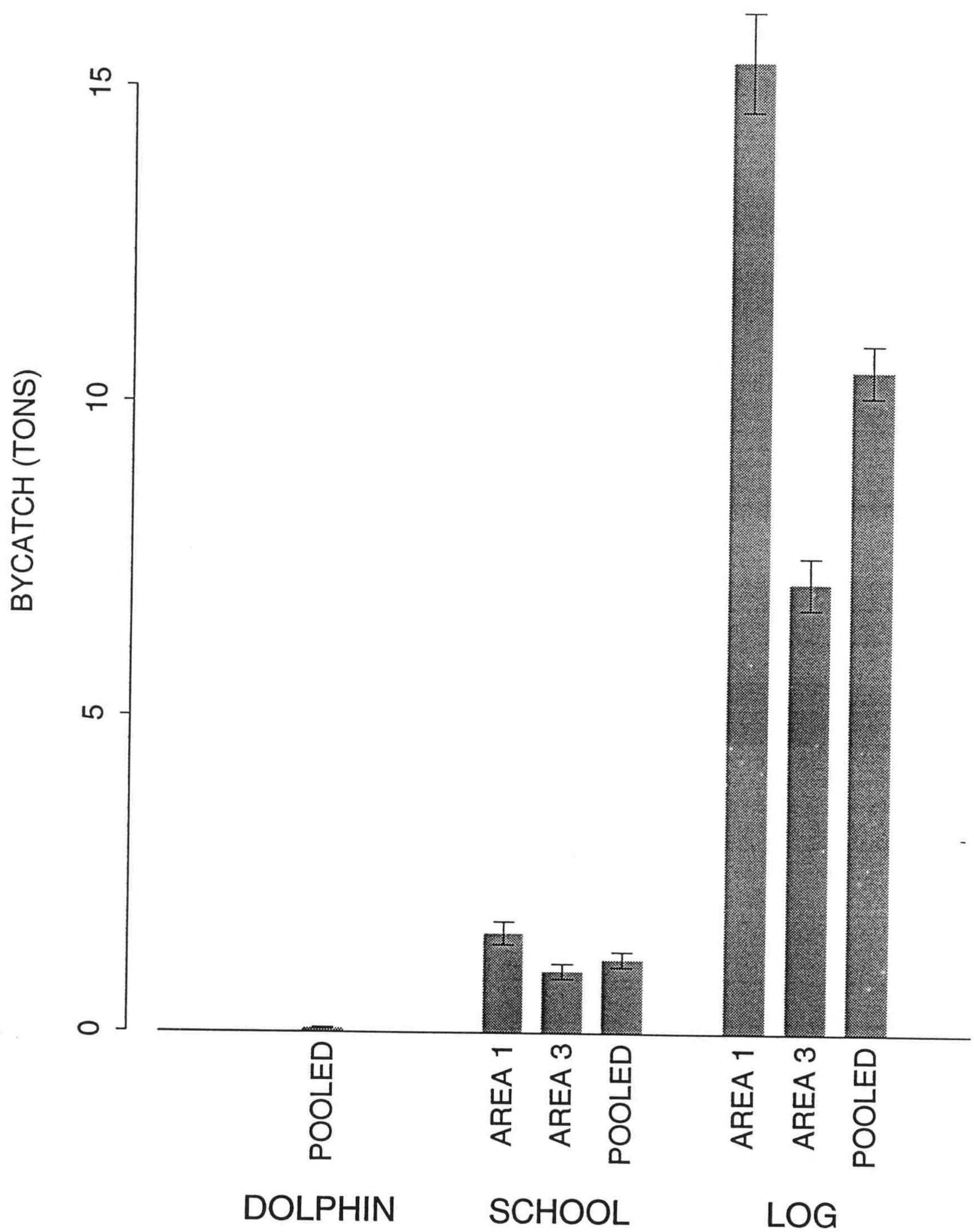

FIGURE 9. Coefficients of variation for estimates of the model parameters (p), (a), and (μ), and for estimates of mean bycatch per set for the U.S. tuna purse-seine fleet fishing in the eastern tropical Pacific Ocean, 1989-1992. Geographic areas as defined in Federal Register (1989). Pooled estimates are fishery-wide, across all areas.

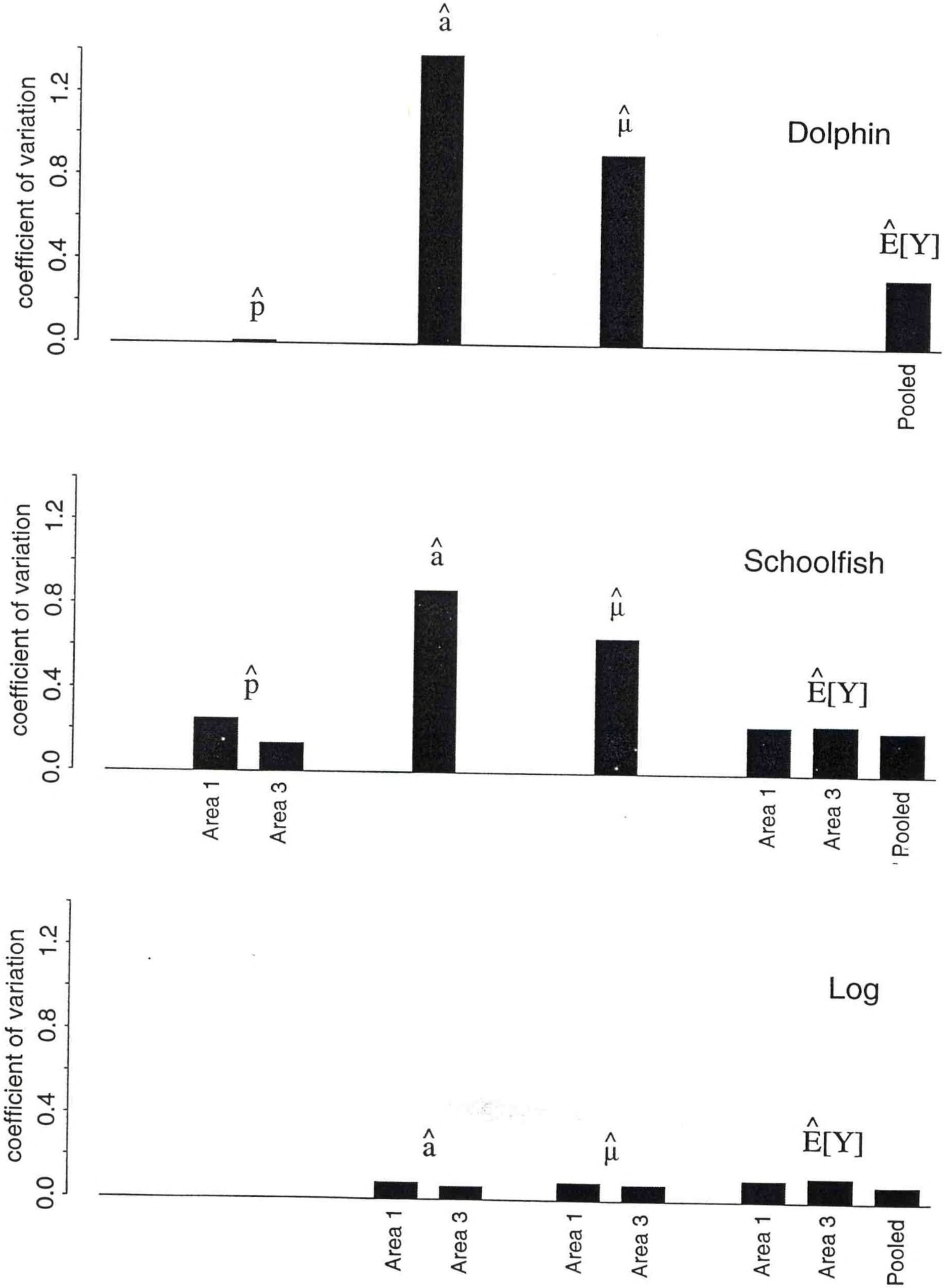

FIGURE 10. Sample histograms of 1000 bootstrap replicates of estimates of the negative binomial parameter (μ), for dolphin, schoolfish, and log sets. See text for a description of the parameter.